N71-22575

NASA TECHNICAL MEMORANDUM

NASA TM X-2264

NASA TM X-2264

CASE FILE COPY

DATA EDITING

by June E. Thompson

Lewis Research Center

Cleveland, Ohio 44135

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION · WASHINGTON, D. C. · APRIL 1971

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM X-2264		
4. Title and Subtitle DATA EDITING		5. Report Date April 1971
		6. Performing Organization Code
7. Author(s) June E. Thompson		8. Performing Organization Report No. E-6060
9. Performing Organization Name and Address Lewis Research Center National Aeronautics and Space Administration Cleveland, Ohio 44135		10. Work Unit No. 129-04
		11. Contract or Grant No.
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D. C. 20546		13. Type of Report and Period Covered Technical Memorandum
		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

An editing system has been developed at Lewis Research Center for a large general time-sharing computer. A dynamic editing display technique and a static microfilm method are described. These editing techniques have been used to reduce the cost of processing large volumes of research data and of eliminating redundant data. They have also been helpful to the engineer in detecting faulty instrumentation and defining processing instructions.

17. Key Words (Suggested by Author(s))		18. Distribution Statement
Research data	Static editing	Unclassified - unlimited
Editing	Display	
Time-sharing	Microfilm	
Dynamic editing		

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	13	$3.00

*For sale by the National Technical Information Service, Springfield, Virginia 22151

DATA EDITING

by June E. Thompson

Lewis Research Center

SUMMARY

An editing system has been developed at Lewis Research Center for a large general time-sharing computer. A dynamic editing display technique and a static microfilm method are described. These editing techniques have been used to reduce the cost of processing large volumes of research data and of eliminating redundant data. They have also been helpful to the engineer in detecting faulty instrumentation and defining processing instructions.

INTRODUCTION

The volume of recorded digital research data at Lewis Research Center far exceeds the amount of data which can be reasonably processed by the existing computer systems. Increasing the computer processing capability would result in a direct increase in cost for additional computing equipment and computing personnel. Since only a portion of the recorded data is essential to the solution of an experimental problem, the processing of redundant or superfluous data can be eliminated without adversely effecting the results of the experiment. In any data processing system identifying and eliminating redundant and faulty data as early as possible in the processing sequence produces the largest reduction in cost. Also, the technique of visually editing the data before processing provides a method of identifying redundant or faulty data and provides the controls to eliminate its processing.

Prior to the work described herein, editing at Lewis was done on a small digital computer with a cathode-ray tube display device and light pen (ref. 1). This digital display was moved from the small computer and interfaced with a large general purpose time-sharing computer. Function keys were added as communication devices to the display system. The display software system, which was rewritten to run on the time-sharing system, utilizes the large core and disc capabilities of the computer. This revision resulted in an expanded dynamic display system on a large time-sharing computer providing user communication by means of function keys, a light pen, and a communication terminal. A second and different approach to editing, using a microfilm recorder, was also developed for the time-sharing system. This technique produces a permanent record of the data on microfilm which may be reviewed at the user's convenience.

The final result is a comprehensive editing system on a time-sharing computer which provides flexability in the editing of research data. The research engineer has the choice of selecting the method which best satisfies his requirements for successfully editing his data. The data editing system described herein has been in operation for several years at Lewis.

Overview of the System

Lewis Research Center is engaged in the recording and processing of large amounts of experimental data. Transient or high speed recording systems located at Lewis or the Plum Brook Station digitize the data at rates from 2 to 30 kilohertz and record them on magnetic tape. A computerized data reduction system, which is available at Lewis, processes and analyzes the data. The data recorded on the tape are the output of test instruments used during the running of an experiment. These data are separated into readings, where a reading corresponds to the data recorded under a unique set of test conditions and identified by a unique number. A reading consists of many data blocks or scans through all channels of a sampling sequence (order in which the channels are digitized). Each channel represents the output of one instrument of the experiment.

The editing system is designed to operate on the IBM 360 Model 67 computer under the control of the TSS/360 time-sharing system. This computer controls various peripheral devices that are functionally available to users wishing to edit recorded data. Since these recorded data are stored on magnetic tape, the computer uses tape drives as its basic input devices for editing. Communication between the display systems and the computer is initiated by any of a number of communications terminals attached to the 360/67. The display of data is done either on a CDC Model DD-20 display device or on a CDC Model DD-280 display.

The DD-20 display consists of a large (15-in. (38.1-cm) diameter) cathode-ray display tube mounted in a console unit and an associated control unit. The control unit contains a separate core memory which stores the control information and data necessary to continuously refresh the display tube. A light pen attached to the display console is used to select points on the screen. Information concerning these points is transmitted to the editing program through the control unit. Function keys on the face of the console are another means of communication between the editor and the display program; these keys permit unique predetermined functions to be requested of the display program. Depressing a function key interrupts the computer and provides identification of the key depressed to the display program. The DD-280 display consists of a 14 by 14 inch (35. 6 by 35. 6 cm) cathode-ray screen and two DD-280 35-millimeter microfilm recorders. The microfilm cameras photograph the plotted information displayed on the face of 5. 25-inch (13. 34-cm) cathode-ray tubes. Data editing may then be performed either dynamically on the screen of the DD-20 display or by using permanently recorded microfilm plots.

The time-sharing system (TSS/360) is a very comprehensive computer operating system which permits many users to simultaneously perform scientific calculations. In other words, one display user may be performing editing tasks at the same time many other computer users are accomplishing tasks of their own. The time-sharing system operates in such a manner that each user seemingly has complete use of the computer.

Dynamic Editing

The dynamic editing system enables the engineer to interact with a digital display of his research data. The display program uses a light pen, function keys, and a communication terminal to facilitate this interaction (fig. 1). Commands to the display program are entered through either the function keys or the terminal. The function keys are a more efficient communications device than the typewriter for requests that do not require descriptive information through the typewriter.

The editor must use the terminal to request the initial reading he wishes to review. Subsequent readings on the tape may be requested for display by using a function key or the terminal. When a requested reading is located, it is read from the tape and stored in the computer. The identifying information and the total number of data blocks in that particular reading are displayed as alphanumeric information on the screen. The display of selected data channels may be initiated at this time. To enter a set of channel numbers through the terminal, a function key is depressed which will request the system to initiate the terminal for input. The terminal then indicates to the user that he may proceed to enter the command (in this case, channel number). The data are displayed as data points as functions of a time plot of up to four channels; each plot is on a separate axis and identified by channel number.

The length of a high-speed data reading frequently exceeds the number of points which can be plotted on DD-20 screen. To enable the engineer to review the complete reading, data from this reading may be plotted in either of two modes - reference or expanded. A reference plot selects representative points from the entire reading to produce a general picture of the data. Areas which the engineer wishes to review in greater detail are selected by the light pen and the plot is expanded to include all data points from this starting point (up to the maximum number of points allowed on the screen). The expanded display may now be shifted along the time axis in either direction. This "move mode" uses two of the function keys to control the direction of the shift and the light pen to select the start or end of the new area. This move mode has been found to be useful for editing readings of great length and is a comprehensive method for locating instrument failure or determining which areas are important for processing.

The appropriate function keys may be used to sequence through additional channels from the same reading, to input the next reading for display, and to add or delete a grid. Appendix A lists the complete set of commands available in the display editing program.

Perhaps the most important function of the light pen is its ability to communicate to the program enough information to determine the magnitude and the real time of a selected data point (fig. 2). The point selected by the light pen is blinked on the display screen. The time and magnitude (in percent of full scale) of this data point are displayed in alphanumeric characters at the bottom of the screen. This information can be used to locate marker channel changes or other significant events in the data.

Various other functions are available to handle specific editing needs. One of these is a display of data in a channel as a function of channel rather than the standard channel as a function of time mode. This display format has proved to be of value to those engineers wishing to observe phase relations in related data channels. Several channels may also be plotted together on the same axis (fig. 2). This feature compares similar data to detect faulty instrumentation. Sufficient information may be obtained from the channels containing the faulty data to enable the engineer to correct the instrumentation before further data are recorded. The engineer may also determine the necessary corrective action which should be taken before processing. The processing program may then be instructed to either avoid processing the faulty data altogether or to substitute other useful or theoretical data for the questionable data. Sometimes it is more useful to tag or code out the channels representing the output of the faulty instruments before processing. A code-out is a tag or bit appended to the data word and transmitted (propagated) with the data through all processing steps. The tagged arithmetic system (ref. 3) performs the calculations with the data values and tags the results rather than replacing the tagged value with a preselected value. On a printout the tag is represented by a literal character to the right of the data value.

Facilities which record very short data readings may be interested in displaying the data for several or all of their readings at one time. This type of display is initiated by requesting that the editing program store the entire set as if it were a single reading. During this mode of display all of the basic editing commands are available to the editor. However, when a data point is selected by the light pen, the reading to which the data point belongs is reported rather than the time of the point. The engineer is thus able to obtain quickly and with ease a display of an entire set of readings and at the same time request information about specific readings in the set.

STATIC EDITING

The static editing system produces a permanent record of selected recorded data through the use of the microfilm recorder and Lewis's microfilm plotting package (ref. 2). This editing program reads the recorded data from magnetic tape and constructs time plots of up to eight channels of data similar to the display constructed by the digital display editing system. The resultant picture is transmitted to the microfilm recorder along with the necessary identifying information (facility, program and reading numbers, initial time of the reading, run date, and channel numbers) and the microfilm is then produced (figs. 3 and 4). The number of time points in a reading may exceed the number of points which can be plotted on a standard frame of microfilm. To produce a continuous plot of the data for the channels requested, the frame size is expanded (called frame butting) and all data points from these channels are displayed regardless of the length of the reading. Data which are questionable, either because they are out of range or are coded out on the input data tape, are flagged (marked). Figure 4 is an example of a data plot which includes faulty or out-of-range data.

The engineer has the option of specifying beforehand that all or only certain channels are to be microfilmed. The user may also request that other specified channels (such as marker channels) are to be scanned for marked changes in the magnitude of the data. The specific times of these changes (in addition to their value if requested) are printed on the microfilm as alphanumeric characters (fig. 5).

Coordinate grids at fixed time intervals may be requested, thereby providing a time scale for the engineer to use in his analysis of the data (fig. 6). Time can then be used as a control to selecting data for processing. Figure 6 illustrates the technique of expanding the Y-scale, which allows the detailed study of, or comparison of, particular measuring instruments.

Data recorded on the high-speed digital recording system are microfilmed when the data are received in the computing center, unless otherwise instructed by the user. The processed film is given to the engineer who may then review his data at any of several microfilm viewers located around the laboratory. The obvious advantage to this method of editing data is that the engineer may review his data quickly and at his convenience. For the engineer who needs only to input times of events to the processing program to reduce the volume of data processed this is a simple, but effective, editing method. If further analysis is required, the engineer may then review his data using the dynamic editing mode.

CONCLUDING REMARKS

The high cost of processing a large volume of research data may be cut considerably by the proper editing of the data. To accomplish this, editing techniques which are flexible enough to satisfy the varying needs of the research engineer, while remaining simple in use, must be made available.

An editing system has been designed to utilize a general time-sharing computer with a large storage capacity. Display devices are accessible through the time-sharing system, and this capability has made possible the development of two editing techniques. A dynamic editing method enables the engineer to interact with his data as displayed in a time-plot mode on a digital display device, whereas the static editing technique is designed to produce a microfilm of the data which may be reviewed at the user's convenience. The engineer has the choice of selecting the method that best satisfies his requirements for successfully editing his data.

The static editing technique provides a way of obtaining a fast overview of the data. Faulty data and instrument failures can be detected and the times of preselected events can be determined. A permanent microfilm record of the data is provided for the engineer. The engineer can review these data at his convenience at a number of viewing stations. The microfilm results are fixed, and further interaction with the data requires dynamic use of the editing program.

The dynamic editing technique allows interaction between the user and his data. The light pen permits the user to dynamically interrogate his data to obtain the time and magnitude. This method is better suited for following up on unexpected events in the data. Since there is only a single dynamic display device, editing time must be scheduled in advance.

A significant reduction in processing costs has resulted with the use of these editing techniques. The engineer can select for processing only those areas necessary to the solution of his problem. This eliminates the cost of processing redundant or nonessential data. Faulty or questionable data can be located and sufficient information obtained to determine the necessary corrective action which should be taken before processing. The processing program may then be instructed to either substitute other useful data for the questionable data, tag or code out the data, or avoid processing data altogether. Reprocessing because of undetected faulty data is then avoided.

Editing is also used by the engineer to detect failing instrumentation and then to correct the problem before further data are recorded. It is also used in analyzing the noise content of the data and in selecting processing techniques to minimize the effects of this noise.

Lewis Research Center,
National Aeronautics and Space Administration,
Cleveland, Ohio, January 7, 1971,
129-04.

APPENDIX - DISPLAY EDITING COMMANDS

The following is a list of the commands available to the user through the terminal or the function keys. Those commands which require descriptive input must be entered through the terminal.

Command	Function
CHAN N1, . . ., N4	initiates the display of channels
COMP N1, . . ., N4	plots up to 4 channels on same axis
EXPD	displays an area selected by the light pen
FIND	locate requested reading on tape and store data in computer
FINI	end editing session
GRID	produces a 10 by 10 grid on screen
MOVL	moves data on screen to earlier time segment
MOVR	moves data on screen to later time segment
NEXC	transmits next sequential channels to screen
NEXR	locate and store next sequential reading
NGRD	deletes grid
PLXY N1, N2	channel as function of channel plot
RDNG N1	specifies reading to be located and stored
REFC N1, N2	displays 2 channels to be used as references for selection of areas of interest
TYPE	initiated by a function key only to allow input to be entered through the terminal

REFERENCES

1. Kronen, Stanley T.: Direct Editing of Research Data by Application of Computer Display Techniques. Presented at the National Telemetering Conference, Boston, Mass., May 10-12, 1966.

2. Kannenberg, Robert G.: Cinematic-Fortran Subprograms for Automatic Computer Microfilm Plotting. NASA TM X-1866, 1969.

3. Bettinger, Paula J.; Manos, Andrew M.; and Armstead, Betty Jo: Tagged Arithmetic. NASA TN D-5370, 1969.

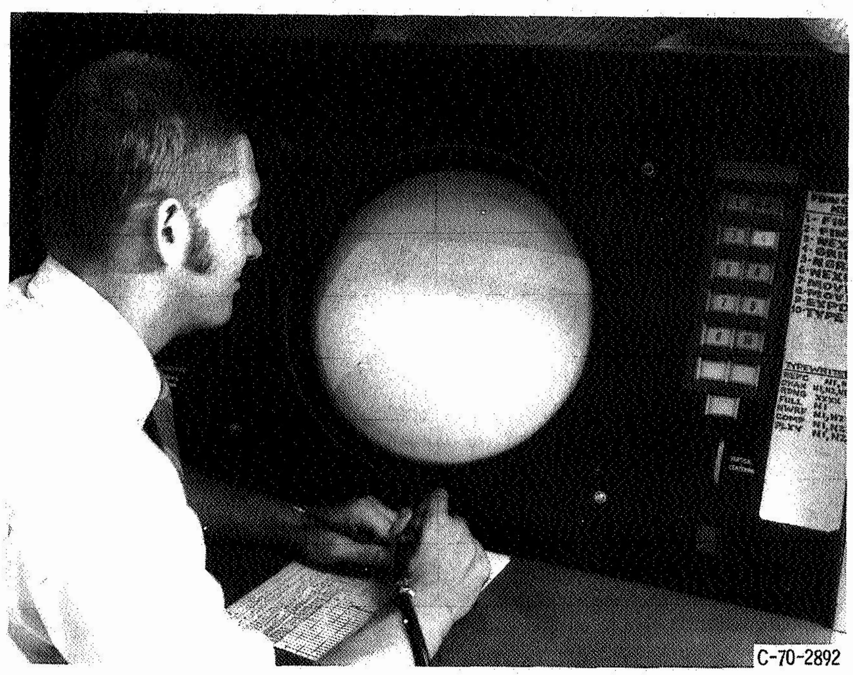
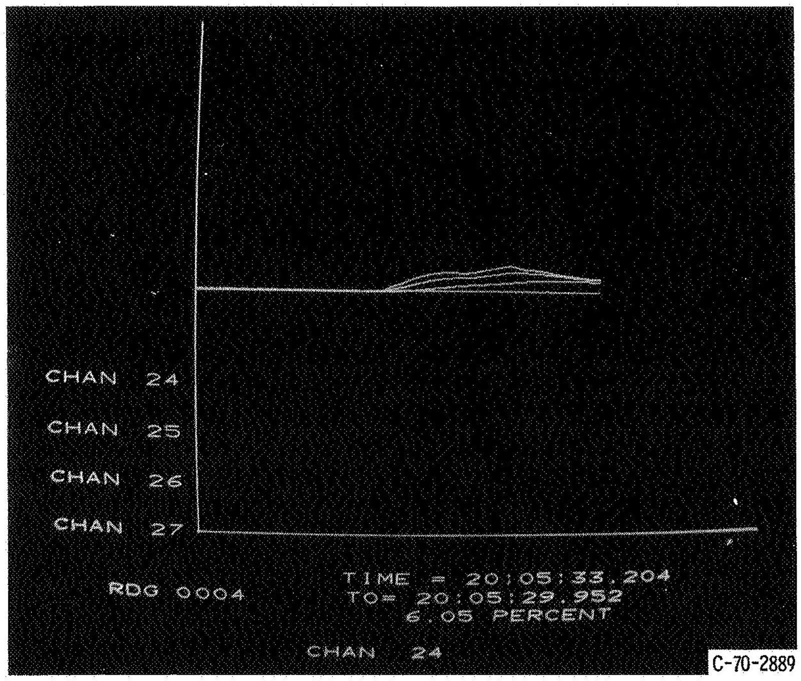


Figure 1. - Communication terminal.



C-70-2889

Figure 2. - Typical data display.

FACILITY	S-40-1
PROGRAM	T017
READING	0001
RUN DATE	06/22/69
INITIAL TIME	18:46:05.918

Figure 3. - Identifying information for recorded data.

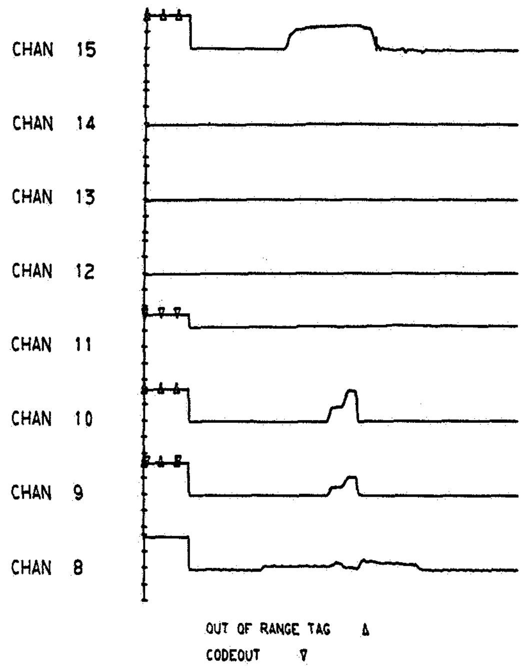


Figure 4. - Microfilm prints of recorded data.

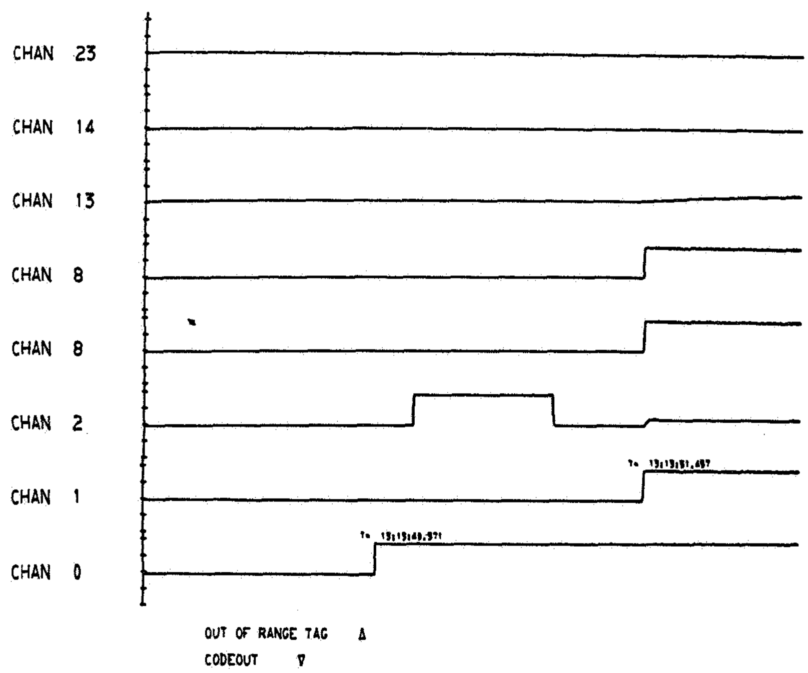
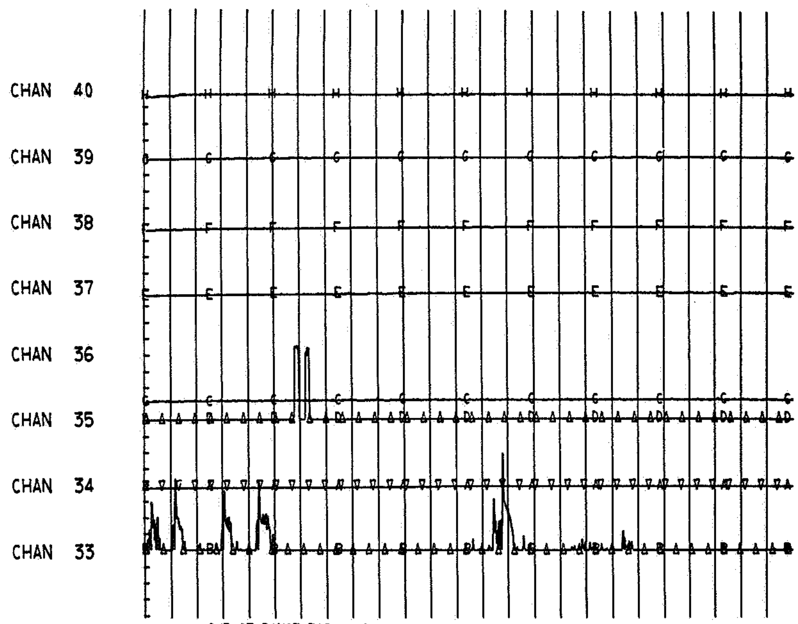


Figure 5. - Marker channel data display.



OUT OF RANGE TAG Δ

CODEOUT ∇

Figure 6. - Time scale added to data display.

NASA-Langley, 1971 —— 8 E-6060

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON, D.C. 20546

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE $300

FIRST CLASS MAIL

POSTAGE AND FEES PAID
NATIONAL AERONAUTICS AND
SPACE ADMINISTRATION

POSTMASTER: If Undeliverable (Section 158 Postal Manual) Do Not Return

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION OFFICE

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546